This is the open access version of the paper:

"Studying Mobile Internet Technologies with Agent Based Mean-Field Models"

LECTURE NOTES IN COMPUTER SCIENCE –VOL 7984 Pages 112-126

DOI 10.1007/978-3-642-39408-9

Studying Mobile Internet Technologies with Agent Based Mean-Field Models

Marco Gribaudo[1], Daniele Manini[2], and Carla Chiasserini[3]

[1] Dip. di Elettronica e Informazione, Politecnico di Milano,
via Ponzio 34/5, 20133 Milano, Italy,
gribaudo@elet.polimi.it
[2] Dip. di Informatica, Università di Torino,
Corso Svizzera 185, 10149 Torino, Italy,
manini@di.unito.it
[3] Dip. di Elettronica e Telecomunicazioni, Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 Torino, Italy,
chiasserini@polito.it

Abstract. We analyze next generation cellular networks, offering connectivity to mobile users through LTE as well as WiFi. We develop a framework based on the Markovian agent formalism, which can model several aspects of the system, including the dynamics of user traffic and the allocation of the network radio resources. In particular, through a mean-field solution, we show the ability of our framework to capture the system behavior in flash-crowd scenarios, i.e., when a burst of traffic requests takes place in some parts of the network service area.

1 Introduction

One of the most evident and urgent challenges in the field of communication networks today is coping with the exponential growth of the wireless data traffic: the average smartphone is expected to generate 2.6 GB of traffic per month by 2016, with a global mobile data traffic that is slated to increase 18–fold by that time [4]. To accommodate such high data-traffic loads, new technologies, such as Long-Term Evolution (LTE), have been introduced to increase the capacity of cellular networks.

The fast uptake of mobile data services, however, indicates that these solutions are not sufficient to meet the intense user demand in many high-density settings all over the world. Thus, a new trend, usually referred to as *mobile data offloading*, has emerged. That is, while the cellular infrastructure will continue to provide essential wide-area coverage and support for high-mobility users, it will be complemented with WiFi hotspots, toward which data traffic should be offloaded whenever possible [9, 13].

Such a scenario calls for a new access network architecture, composed of base station units (BSs) that may host several radio interfaces, hence provide Internet connectivity to mobile users through different communication technologies (e.g., LTE and WiFi) [5, 2]. Beside meeting the users demand, this network paradigm

will imply low real-estate and deployment costs, as well as low upgrading and security costs, thus saving the operators millions of dollars in captial and operational expensitures (CAPEX/OPEX).

It is important to stress that wireless communication technologies, like LTE and WiFi, are already mature, and that commercial products implementing the aforementionwd paradigm are available on the market, e.g., [2, 1]. What is missing, however, is the definition of algorithms to make the above access network work efficiently. In particular, the functionalities of the system should be optimized so as to support the huge amount of data that wireless users are expected to consume/generate, while meeting the requirements in terms of quality of service, energy consumption and cost.

In this paper, we focus on the above aspect and develop a framework for the analysis of different policies regulating the Internet connectivity of mobile users. Specifically, we consider that the network service area is covered by a number of BSs, each of them hosting both an LTE and a WiFi radio interface. A user can connect to the Internet through either technologies, provided that enough radio resources are available to serve the user. Typically, widely popular communication devices, such as smartphones, implement a simple connectivity policy, which is known as "WiFi first": a user always connects to a WiFi hotspot when available. The rational behind this choice is that WiFi connectivity is much less costly than the cellular one: through WiFi the per-byte cost of data transfers can be reduced by 70% per one estimate [3]. However, WiFi may offer a much slower data transfer than LTE, especially when several users are accessing the same hotspot, or the radio propagation conditions are not favourable because of the user mobility or the presence of obstacles [12, 15, 5]. This clearly motivates the need for a framework that allows to evaluate various connection policies, beside "WiFi first".

In the following, we make a first step toward the solution of this issues. Specifically, we propose an analytical model and show how mean-field analysis can be successfully applied to the study of the system dynamics. We remark that a similar study carried out through simulation would imply very long computation times, as each communication node in the network (either user or BS) would have to be equipped with two radio interfaces, each of which would require different models representing the signal propagation as well as the node protocol stack.

2 Network scenario and mobile Internet technologies

We consider a urban area (typically characterized by high user density) and covered by an LTE cellular network. The network system is composed of several BSs, each of them covering an area that we will be referred as *coverage area*. Colocated with the LTE interface, there is a WiFi radio (IEEE 802.11a/g/n), so as to implement a hostspot whose coverage coincides with that provided by the LTE technology.

We only consider data transfers, such as content dowloading or video streaming, as voice calls cannot be supported through LTE. Also, we focus on downlink data transfers (from the BS to the users) since traffic is typically highly asymmetric, with a large amount of data flowing from the Internet towards the users. Clearly, users that are under the coverage of more than one BS, can access any of them, although with different quality of service. In particular, it is fair to assume that the link quality increases as the distance between a user and a BS decreases, and that the better the link quality, the higher the transmission rate that the link end points can use. Furthermore, we consider that the users will not significantly move while receiving a data transfer from the Internet.

For clarity, before presenting the model we developed, we summarize the main characteristics of the two technologies that are at the basis of our analysis.

WiFi. The WiFi technology is specified by the IEEE 802.11 standard. It allows for the implementation of a wireless local network (WLAN) or hotspot, composed of an access point (AP) and user communication devices. Users that are connected to a hotspot can download/upload traffic from/to the Internet, or exchange traffic between each other; all traffic however is handled by the AP. In order to transmit data over the wireless channel, users and AP employ a totally distributed scheme, namely, a carrier sense multiple access (CSMA) technique enhanced with a collision avoidance mechnism. In a nutshell, whenever a network node wishes to transfer data, it senses the channel. If idle, it will transmit; otherwise, it defers its transmission by a random time. Clearly, the higher the traffic load within the WLAN, the higher the data latency and the probability that two or more transmissions start at the same time and, thus, fail. It is easy to show that, on the long run, such a channel access scheme provides equal access opportunities to all nodes (both users and AP).

Furthermore, it is important to note that at the physical layer the nodes may use different transmission speeds (i.e., data rates), depending on the propagation conditions between the sender and the intended receiver: the higher the packet error probability that a node experiences, the lower the rate it uses to transmit its data. As an example, in the "a" version of the 802.11 standard 8 values of data rate are possible, ranging from 6 up to 54 Mb/s. It follows that the transmission time of the data packets will depend on both the packet size and the rate used by the sender to transmit toward the intended recepient. This aspect may have severe consequences on the throughput experienced by the users connected to the same hotspot. Indeed, as noted in [14], "slower" senders occupy the channel longer preventing others from transmitting. Since the WiFi technology provides equal channel access opportunities to all nodes, as the number of slow senders increases, the fast ones will be able to transmit less often thus experiencing a low throughput even if they can employ a high data rate.

Finally, we remark that different frequency channels can be used by the WiFi technology. Again, with reference to the 802.11a standard, 8 disjoint channels around 2.4 GHz are available, and each hotspot can select one among these frequency bands. Typically, nearby hotspots employ different channels so as to avoid co-channel interference.

LTE. The LTE cellular technology is marketed as 4G and is already available in several regions, such as Australia, North America, and Scandinavian countries. Thanks to the use of multiple antennas at the transmitter and receiver, it implements multiple-input-multiple-output (MIMO) trasnmission techniques, i.e., the simultaneous transmissions of a number of flows as high as the number of antennas available at the communication end points. The transmission speed is therefore significantly higher than in the previous cellular technologies, and, in principle, it can reach up to 300 Mb/s in downlink (from the BS to the users) and 50 Mb/s in uplink (from user to BS). Current implementations, however, allow for much lower data rates, leading to a throughput of up to 30 Mb/s in downlink and of about 14 Mb/s in uplink.

As in most of the previous cellular technologies, a frequency division duplex (FDD) technique is used for data transmission: two separate frequency channels, say of 1.4 MHz each, are used for uplink and downlink traffic, respectively. Also, within an area covered by a BS, the access to the uplink channel is controlled by the BS itself, which schedules the user transmissions. Focusing on the downlink direction, an orthogonal frequency division multiple access (OFDMA) technique is employed, jointly with a time division scheme. Indeed, the downlink frequency channel is divided into several narrow-band subchannels. Disjoint subsets of such channels are then used by the BS to simultaneously transmit towards different users. Each subchannel is 15-MHz wide and each subset is composed of 12 subchannels. Time is divided into frames that are 10 ms long; each of them is further divided into 10 subframes. In the following, we consider the usage of a subset of 12 subchannels for a 1-ms duration (i.e., a subframe) as the granularity used by the BS to allocate radio resources to a data flow, i.e., the so-called Physical Radio Block (PRB). We remark that the BS may allocate one or more PRBs to transmit at the same time towards the same user. Also, as in the case of the WiFi APs, LTE interfaces at neighboring BSs should use different frequency channels so as to avoid interference.

3 Modelling LTE and WiFi technologies

As outlined above, the system under study is composed of two main types of communication nodes, interacting with each other: users and BSs. The spatial distribution of such nodes plays a critical role, as a user can access a BS only if its position is within the coverage area of that BS. These peculiariries of the system prevent the use of standard state-space modeling techniques, such as Queueing Networks, Stochastic Petri Net, or Process Algebras, since all of them would suffer from the problem of state-space explosion.

In this work, we therefore resort to a technique based on the Markovian Agent formalism [10], and exploits the Mean Field Analysis [6, 7]. In particular, our approach leverages the methodology presented in [8] to compute the model solution.

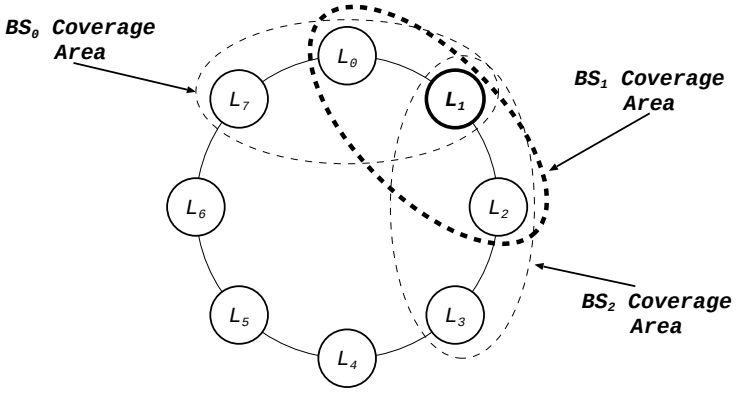


Fig. 1. An example of the ring topology under study.

3.1 Network topology and model

To simplify the presentation, we consider a network with ring topology, composed of N geographical areas, hereinafter called locations and denoted by $L_0, \ldots, L_{N-1}$. Given the topology, each location L_i has two neighboring locations, L_{i-1} and L_{i+1}, where the generic pedex i has to be intended as $i \bmod N$. At the center of each location L_i, there is a BS, denoted by BS_i ($i = 0, \ldots, N-1$), whose coverage area extends over three locations, namely, L_i, L_{i-1} and L_{i+1}. It follows that users located within L_i can access BS_i, or equivalently, BS_{i-1} and BS_{i+1}. Fig. 1 shows an example of the ring topology used in our model, for $N = 8$. In particular, circles represent the locations $L_0, \ldots, L_7$, while each dotted ellipse corresponds to the coverage area of a BS.

We call *agent* a portion of the model that describes the behavior of the communication nodes within a given location L_i. Specifically, we talk about user or BS agent depending on whether an agent refers to the behavhior of the users or of the BS. We then define as *access group* i (also denoted by AG_i) the the set of users that access the network using BS_i.

All users agents have exactly the same structure and include the same states, as shown in the left side of Fig. 2. Following the definition given in [8], the agents in the different locations correspond to different agent classes, all deriving from the same metaclass. As depicted in the figure, each agent is composed of 7 states. The first six account for the particular technology that a user is using to access the Internet (namely, LTE or WiFi), while the last state is used to count the user connection attempts that fail due to the unavailability of radio resources. The location from which the technology is accessed is denoted by the following labels: *local* for the BS in the location the user agent refers to, and *cw* and *ccw* for the BS in, respectively, the clockwise and the counter-clockwise neighboring location. In particular, we consider BS_{i+1} and BS_{i-1} being, respectively, the clockwise and counter-clockwise neighbors of BS_i. Finally, the label *Loss* denotes the state counting the service requests that have not been accomodated. The interaction

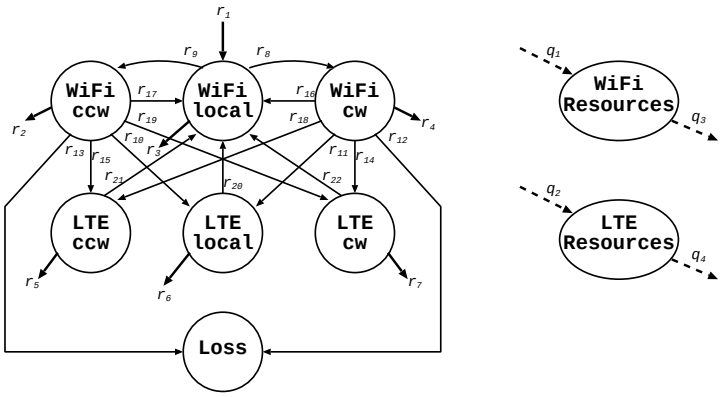


Fig. 2. Mean-field models of a user agent (left) and of a BS agent (right).

among the users is determined by the transition rates, which are functions of the number of users in the each state of the local and of the neighboring locations. Since the transition rates are the key element of the model, they will be described in depth in Section 3.2.

Similarly to the user agent, we define another metaclass to describe the BS status. Again, each BS agent corresponds to a class deriving from the same metaclass. Thus, all BS agents have the same struture: they include two states, each counting the number of allocated radio resources, one for WiFi and the other for LTE (as illustrated in the right side of Fig. 2). Note that by radio resources we mean frequency channels in WiFi and PRBs in LTE. The incoming/outcoming arcs account for the dynamic allocation of BS radio resources.

Fig. 3 further clarifies the relationship among user agents, locations, and access groups. Recall that each access group AG_i ($i = 0, 1, 2$ in the figure) refers to the set of users accessing BS_i, while each agent includes the possible states taken by users located within L_i. For instance, AG_1 includes states that belong to three different agents (each denoted by a different color), as BS_1 can be accessed by users that are in L_0, L_1 and L_2. Thus, AG_1 contains those states of the three user agents that correspond to the usage of radio resource handled by BS_1 (e.g., for the agent referring to L_0: WiFi *cw* and LTE *cw*).

3.2 Functional rates

The interaction among the agents is determined by the transition rates; in this section we present the functions that allow to compute their values.

With regard to the BS agents, the control of the resource allocation at each BS is described through the rates q_1, q_2, q_3 and q_4. In the following, we will consider a static resource allocation, hence we set the above rates to the same constant value.

The evolution of a user agent through the model states is determined by the values assigned to the following rates (see also Fig. 2):

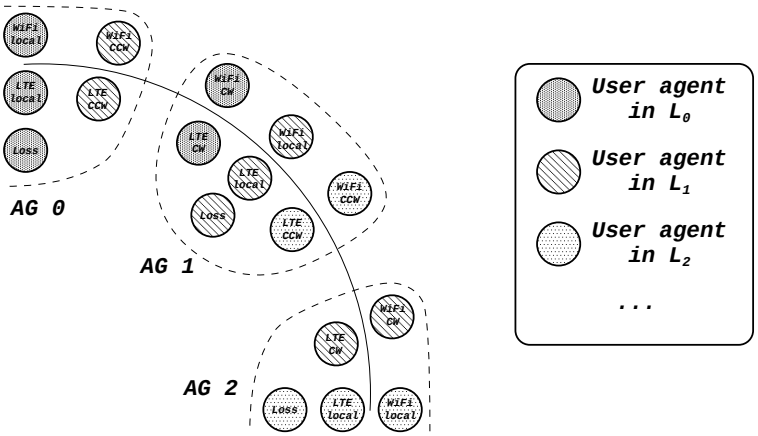


Fig. 3. Relation among user agents, locations, and access groups; different user agents are denoted by different colors.

- Arrival of users requests (r_1);
- Service of requests (r_2, r_3 and r_4 for WiFi; r_5, r_6 and r_7 for LTE);
- WiFi switching from the local BS to a neighboring one (r_8, r_9);
- WiFi switching to LTE (r_{10}, r_{11} from neighboring BSs to the local one; r_{14}, r_{15}, r_{18}, r_{19} for neighbouring BSs);
- Loss of requests (r_{12}, r_{13});
- Upgrading: WiFi switching from a neighboring BS to the local one (r_{16}, r_{17}), and technology switching from LTE to WiFi (r_{20}, r_{21}, r_{22}).

Below, we introduce the main concepts and, in few cases, also the equations that we used to derive the rates. The computation details are reported in Table 1, where the following notations are exploited:
$I_1(condition) = 1$ if *condition* is true, 0 otherwise
$I_2(a, b) = 1$ if $a > b$, 0.5 if $a = b$, 0 if $a < b$ [4]
and where $n^{[i]}_{\tau-j}$ denotes the number of users in $\tau - j$ state of the user agent referring to location L_i, with $j \in \{local, cw, ccw\}$ and $\tau \in \{WiFi, LTE\}$.

Next, we describe all the possible events that trigger a transition between two states of the user agent, and the corresponding rates. Note that for what concern technology switching criteria, we model a specific system behavior based on arbitrary rules which can be easily refined to study more complex dynamics.

[4] To improve the smoothness of the solution, in the experiments we used for $I_2(\cdot)$ the *Sigmod* function defined as follows:

$$I_2(a, b) = \frac{1}{1 + e^{\alpha(b-a)}}$$

with α very large (we used $\alpha = 100$).

Arrivals. We denote by $\lambda_i(t)$ the number of requests per second generated by users in location L_i; note that such parameter has both a spatial and a time dependency. For the generic user agent referring to location L_i, we set $r_1 = \lambda_i(t)$. Clearly, the request rates are not affected by the system status. While deriving our results, we will generate a burst of requests in location subsets, in order to observe the network response.

Services The WiFi service rates, to be assigned to r_2, r_3 and r_4, are derived through the following formula:

$$\sigma^{[i]}_{WiFi} = \frac{\mu_{WiFi} \times n^{[i]}_{WiFi-resources}}{n^{[i]}_{WiFi-local} + \alpha(n^{[i-1]}_{WiFi-cw} + n^{[i+1]}_{WiFi-ccw})} \tag{1}$$

where:

- μ_{WiFi} is the WiFi throughput (connection speed) when there is only one user accessing the interface;
- $n^{[i]}_{WiFi-resources}$ is the number of WiFi resources available at BS_i;
- $n^{[i]}_{WiFi-local}$ is the number of users in L_i accessing the WiFi interface at BS_i;
- $n^{[i-1]}_{WiFi-cw}$ is the number of users in location L_{i-1} that access the WiFi interface at BS_i, i.e., at their clockwise (cw) BS;
- $n^{[i+1]}_{WiFi-ccw}$ is the number of users in location L_{i+1} that access the WiFi interface at BS_i, i.e., at their counter-clockwise (ccw) BS;
- $\alpha \geq 1$ is a factor taking into account the throughput reduction due to far-away users accessing the WiFi interface at BS_i; such far-away users are those in locations L_{i-1} and L_{i+1}.

In particular, we have: $r_2 = \sigma^{[i-1]}_{WiFi}$, $r_3 = \sigma^{[i]}_{WiFi}$ and $r_4 = \sigma^{[i+1]}_{WiFi}$.We remark that the factor α is necessary in order to account for the anomaly effect typical of the WiFi technology, as mentioned in Section 2.

Let us call $\sigma^{[i]}_{LTE}(j)$ the LTE service rate assigned to a user in L_j accessing BS_i. For $i = 0, \ldots, n-1$, we have $r_5 = \sigma^{[i-1]}_{LTE}(i)$, $r_6 = \sigma^{[i]}_{LTE}(i)$ and $r_7 = \sigma^{[i+1]}_{LTE}(i)$. The LTE service rate, can be defined as:

$$\sigma^{[i]}_{LTE}(j) = \frac{\mu^{i,j}_{LTE} \times n^{[i]}_{LTE-resources}}{n^{[i]}_{LTE-local} + n^{[i-1]}_{LTE-cw} + n^{[i+1]}_{LTE-ccw}} \tag{2}$$

where:

- $\mu^{i,j}_{LTE}$ is the LTE throughput (connection speed) corresponding to one radio resource assigned from BS_i to a user accessing from L_j. Note that in LTE a radio resource is assigned exclusively to one user, however, depending on the radio propagation conditions, the user can employ a higher or a lower transmission rate. This motivates the dependency of the parameter on the user location (L_j), as the farther away the user is, the lower its connection speed;

- $n^{[i]}_{LTE-resources}$ is the number of LTE channels at the local BS (BS_i);
- $n^{[i]}_{LTE-local}$ is the number of users in L_i accessing the LTE interface at BS_i;
- $n^{[i-1]}_{LTE-cw}$ is the number of users in location L_{i-1} that access the LTE interface at BS_i, i.e., at their clockwise (cw) BS;
- $n^{[i+1]}_{LTE-ccw}$ is the number of users in location L_{i+1} that access the LTE interface at BS_i, i.e., at their counter-clockwise (ccw) BS.

WiFi switching to neighbouring BSs. The switching from local to a neighbouring WiFi interface is forced if the local service rate r_3 drops below the minimum threshold (μ_{min}). The neighbouring BS with the lowest load is selected (r_8, r_9), if the two locations have the same load the destination is randomly selected between them. The latency due to the interface switching is denoted by $1/\mu_{BS-Sw}$.

Switching from neighboring WiFi to local LTE. When the WiFi service rate in neighboring BSs drops below a minimum threshold (μ_{min}), the users attempt to migrate to the LTE interface of the local BS (transitions with associated rates r_{10} and r_{11}). Indeed, although connectivity through LTE is more costly, we expect that users are willing to pay a higher price provided that they can obtain a sufficient connection speed. The transition success depends on the actual load of the local LTE interface; in particular, the following condition must hold:

$$n^{[i]}_{LTE-local} + n^{[i-1]}_{LTE-cw} + n^{[i+1]}_{LTE-ccw} < n^{[i]}_{LTE-resources} \times B \times \gamma_{min} \qquad (3)$$

where $n^{[i]}_{LTE-local}$, $n^{[i-1]}_{LTE-cw}$, $n^{[i+1]}_{LTE-ccw}$ and $n^{[i]}_{LTE-resources}$ are defined as above, while B is number of 1-ms time intervals (subframes) for each LTE frequency subset (see Section 2), and γ_{min} is the utilization threshold of LTE resources at a BS. The user migration takes place (with a latency due to the to technology switch, $1/\mu_{\tau-Sw}$) if the utilization level of the LTE resources at the target BS is under γ_{min}. This avoids overloading the LTE interface and, thus, the need to move traffic flows back again to neighboring WiFi interfaces after some time.

Switching from WiFi to LTE at neighboring BSs. The switch from neighboring WiFi to local LTE can fail if condition (3) is false. In this case, users may try to move towards an LTE interface at a neighboring BS. The switch is successful if either one of the neighboring BSs exhibits a utilization level of their LTE resources lower than γ_{min}. If both BSs can accomodate the user, the rate function selects the BS with the lower LTE load. In case of a tie, the destination BS is randomly selected. At last, we recall that this algorithm is used for deriving r_{14}, r_{15}, r_{18} and r_{19}.

Loss of requests. A user data request fails when it cannot be accomodated at any BSs (using either LTE or WiFi), i.e., there are no resources available.

Specifically, when the WiFi service rate is below the μ_{min} threshold, the function for rates r_{12} and r_{13} verify the condition (3) for the LTE interface at the local as well as at the neighboring BSs. If the condition is always false, the request cannot be accomodated and the number of request failures is incremented.

Upgrade. The upgrading to the local WiFi interface can take place from that at neighbouring BSs as well as from the LTE interface at either local or neighbouring BSs. Specifically, when rate r_3 is beyond a given threshold (μ_{max}), the users can migrate to the local WiFi interface. Note that such a transition indeed represents an upgrade since using the (non-overloaded) WiFi inteface at the local BS implies high speed connectivity at low cost. The time required by the transition depends on the type of switch that is performed, i.e., the technology does not change but the BS does ($1/\mu_{BS-Sw}$ for r_{16} and r_{17}) or the technology (and possibly the BS) changes ($1/\mu_{\tau-Sw}$ for r_{20}, r_{21} and r_{22}).

4 Analysis and numerical results

The model proposed in Section 3 is solved using the techniques described in [8]. In particular, the model of the agent presented in Fig. 2 is used to determine, for each location L_i, two matrices and a vector: the transition matrix $\mathbf{C}^{[i]}$, the death matrix $\mathbf{D}^{[i]}$ and the birth vector $\mathbf{b}^{[i]}$. The resulting matrices and vectors are as follows:

$$\mathbf{C}^{[i]} = \begin{vmatrix} -() & r_{17} & 0 & r_{15} & r_{10} & r_{19} & r_{13} \\ r_9 & -() & r_8 & 0 & 0 & 0 & 0 \\ 0 & r_{16} & -() & r_{18} & r_{11} & r_{14} & r_{12} \\ 0 & r_{21} & 0 & -() & 0 & 0 & 0 \\ 0 & r_{20} & 0 & 0 & -() & 0 & 0 \\ 0 & r_{22} & 0 & 0 & & -() & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix} \quad \mathbf{D}^{[i]} = diag \begin{vmatrix} r_4 \\ r_3 \\ r_2 \\ r_5 \\ r_6 \\ r_7 \\ 0 \end{vmatrix} \quad \mathbf{b}^{[i]} = \begin{vmatrix} r_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}^T \tag{4}$$

Matrix $\mathbf{C}^{[i]}$ must be an infinitesimal generator: all its rows must sum up to zero. To simplify the presentation, we have used the notation "$-()$" to identify the sum of the other elements in the row, changed of sign. The row represents, respectively, the following states of the agent: $WiFi-ccw$, $WiFi-local$, $WiFi-cw$, $LTE-ccw$, $LTE-local$, $LTE-cw$, $Loss$. Note that, since the $Loss$ state is absorbing, the corresponding row is zero. The number of users in a given state is collected in a row vector $\mathbf{n}^{[i]} = |n^{[i]}_{WiFi-ccw}, \ldots, n^{[i]}_{Loss}|$, and the evolution of the system is computed by solving the following equations[5], one for each location L_i ($i = 0, \ldots, n-1$):

$$\frac{d\mathbf{n}^{[i]}}{dt} = \mathbf{n}^{[i]} \left(\mathbf{C}^{[i]} - \mathbf{D}^{[i]} \right) + \mathbf{b}^{[i]}. \tag{5}$$

[5] Note that all vectors and matrices depend on time. However, we have omitted the time dependency in order to simplify the notation.

Table 1. Rate functions

Rates	Functions
r_1	$\lambda_i(t)$
r_2	$\sigma_{WiFi}^{[i-1]}$
r_3	$\sigma_{WiFi}^{[i]}$
r_4	$\sigma_{WiFi}^{[i+1]}$
r_5	$\sigma_{LTE}^{[i-1]}(i)$
r_6	$\sigma_{LTE}^{[i]}(i)$
r_7	$\sigma_{LTE}^{[i+1]}(i)$
r_8	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i]} < \mu_{min}) \times$ $I_2(\sigma_{WiFi-local}^{[i+1]}, sigma_{WiFi-local}^{[i-1]})$
r_9	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i]} < \mu_{min}) \times$ $I_2(\sigma_{WiFi-local}^{[i-1]}, sigma_{WiFi-local}^{[i+1]})$
r_{10}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i-1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} < n_{LTE-resources}^{[i]} \times B \times \gamma_{min})$
r_{11}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i+1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} < n_{LTE-resources}^{[i]} \times B \times \gamma_{min})$
r_{12}	$\mu_{fail} \times I_1(\sigma_{WiFi}^{[i+1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_1(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]} \geq n_{LTE-resources}^{[i+1]} \times B \times \gamma_{min}) \times$ $I_1(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]} \geq n_{LTE-resources}^{[i-1]} \times B \times \gamma_{min}) \times$
r_{13}	$\mu_{fail} \times I_1(\sigma_{WiFi}^{[i-1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_1(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]} \geq n_{LTE-resources}^{[i+1]} \times B \times \gamma_{min}) \times$ $I_1(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]} \geq n_{LTE-resources}^{[i-1]} \times B \times \gamma_{min}) \times$
r_{14}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i+1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_2(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]}, n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]}) \times$ $I_1(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]} < n_{LTE-resources}^{[i+1]} \times B \times \gamma_{min})$
r_{15}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i-1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_2(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]}, n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]} \times$ $I_1(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]} < n_{LTE-resources}^{[i-1]} \times B \times \gamma_{min})$
r_{16}, r_{17}	$\mu_{BS-Sw} \times I_1(\sigma_{WiFi}^{[i]} > \mu_{max})$
r_{18}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i+1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_2(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]}, n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]}) \times$ $I_1(n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]} < n_{LTE-resources}^{[i-1]} \times B \times \gamma_{min})$
r_{19}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i-1]} < \mu_{min}) \times$ $I_1(n_{LTE-local}^{[i]} + n_{LTE-cw}^{[i-1]} + n_{LTE-ccw}^{[i+1]} \geq n_{LTE-resources}^{[i]} \times B \times \gamma_{min}) \times$ $I_2(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]}), n_{LTE-local}^{[i-1]} + n_{LTE-ccw}^{[i]} + n_{LTE-cw}^{[i-2]}) \times$ $I_1(n_{LTE-local}^{[i+1]} + n_{LTE-cw}^{[i]} + n_{LTE-ccw}^{[i+2]} < n_{LTE-resources}^{[i+1]} \times B \times \gamma_{min})$
r_{20}, r_{21}, r_{22}	$\mu_{\tau-Sw} \times I_1(\sigma_{WiFi}^{[i]} > \mu_{max})$

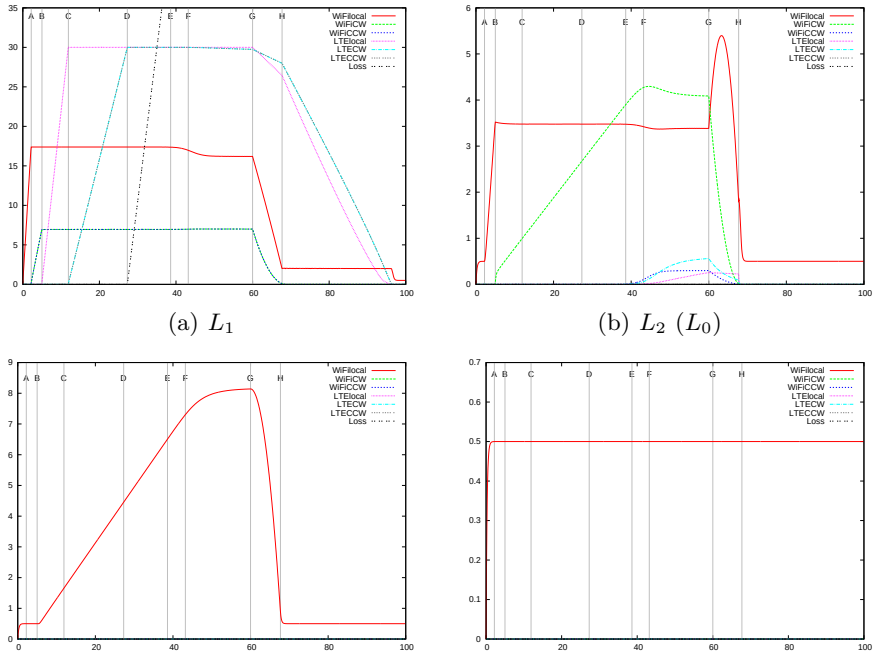
Equation (5) can be solved using a suitable numerical algorithm. In our work, we have used the *Runge-Kutta with adaptive step-size control* discretization method [16].
Note that some rates include indicator functions that can cause problems to the mean field technique. In current experiments we found that the introduction of sigmoid functions instead of indicators is enough to obtain accurate results. However more general approaches based on higher moment approximations such as [11] can provide better results at the expense of greater number of equations.

4.1 Results: Scenario with request burst in one BS

Here, we show the ability of the proposed framework to capture the system dynamics. We consider the topology in Fig. 1 and represent a flash crowd traffic scenario in location L_1. To this end, we set $\lambda_i(t) = 2$ requests/s, $\forall t \in [0, 100]$, for $i \neq 1$, and $\lambda_1(t) = 12$ when $t \in [0, 60)$ and $\lambda_1(t) = 2$ when $t \in [60, 100]$; this corresponds to a request burst affecting L_1 for 60% of the observation time. Recall that the radio resource allocation at the BSs is static, and that all BSs have the same number of WiFi and LTE resources available.

Fig. 4 highlights the impact of the above flash crowd traffic scenario on the load level of the different BSs and on the service provided by the network system. Specifically, the plots depict, for every location, the temporal evolution of the number of user agents in each state. Fig. 4(a) refers to L_1 (i.e., the location affected by the request burst), Fig. 4(b) refers to the neighboring locations L_2 and L_0, Fig. 4(c) represents the behavior of the users in L_3 and L_7, and finally Fig. 4(d) reflects the situation of the users in all the remaining locations. Note that, due to the ring shape of the topology and the considered traffic scenario, the users in locations L_0 and L_2 exhibit the same behavior (except for the fact that *cw* and *ccw* states are inverted); the same observation holds for the pairs (L_3,L_7) and (L_4,L_6), as well as for L_5, L_4 and L_6. Thus, for simplicity, in the following we will refer only to L_1, L_2, L_3 and L_4.

To ease the discussion of the results, in the plots we added vertical lines to indicate the following main events that take place in the system.
Event A: Because of the high traffic load, local WiFi resources at BS_1 saturate (red line in Fig. 4(a)). As a consequence, the local service rate drops below μ_{min} and the users migrate toward the neighboring WiFi intefaces (see the overlapping blue and green lines in Fig. 4(a), and the red line in Fig. 4(b)).
Event B: Also the neighboring WiFi resources (e.g., BS_2) saturate (see, e.g., blue and green lines in Fig. 4(a)), and users in L_1 start accessing the Internet through the local LTE resources (purple line). Also, since at BS_2 all WiFi resources are allocated to users in L_1, the users in L_2 migrate to the neighboring (clock-wise) WiFi interface, i.e., BS_3 (green line in Fig. 4(b), and red line in Fig. 4(c)).
Event C: In L_1 the local LTE saturates (purple line in Fig. 4(a)) and users start using LTE resources at neighboring BSs (light blue and gray lines).
Event D: All resources at BS_1, as well as those in neighboring BSs that can be accessed by users in L_1, are saturated. As a consequence, requests start to be



(a) L_1 (b) L_2 (L_0)

(c) L_3 (L_7) (d) L_4 (L_5, L_6)

Fig. 4. Transient evolution of the user agents in locations $L_0, \ldots, L_7$. Main events are represented by vertical lines.

dropped (black line in Fig. 4(a)). For representation purposes, the curve representing the number of losses is truncated; it actually increases till the request burst in L_1 ends (event G), reaching a total of 135 dropped requests over the whole observation time. We remark that such high losses could be avoided if a dynamic resource allocation were implemented in the system.

Events E: All radio resources at BS_2 are allocated to serve local users and those in L_1; the WiFi resources are saturated at BS_3 too. It follows that users in L_2 start accessing the Internet through LTE at BS_3 (light blue line in Fig. 4(b)); in addition, there are some oscillations in the use of the WiFi resources at BS_1 (blue line in Fig. 4(b)).

Events F: The situation arised in correspondence of event C persists. The WiFi quality of service at BS_3, however, further degrades, hence those users in L_2 that were accessing the BS_3 WiFi resources move back to their local LTE (purple line in Fig. 4(b)).

Events G: The request burst in L_1 ends, thus the traffic load at all BSs start to decrease. BS_2 continues to serve the remaining requests by using both WiFi (whose load increases, see the red line in Fig. 4(b)) and LTE resources.

Events H: The decrease of the traffic load in L_1 allows the handover from LTE (purple, light blue and gray lines, in Fig. 4(a)) to WiFi (red line). Note also that, as the backlog of requests is served by BS_1, the load of the local WiFi resources

(red lines) settles at the same value as that exhbited by the other BSs (namely, 0.5, which is consistent with the generation rate of 2 requests per second).

5 Conclusions and future work

We presented an analytical framework based on the Markovian agent formalism, which models next generation cellular networks. We envisioned a system where base stations can provide Internet connectivity through the LTE as well as the WiFi technology, and we highlighted how our framework can model the different dynamics of the system reflecting the user traffic. By solving the model through a mean-field based methodology, we also showed that the framework can well capture the system behavior in flash-crowd scenarios.

Future work will focus on the model validation through simulation, as well as on the study of dynamic resource allocation strategies and of user connectivity policies. Metrics such as throughput, latency, user energy consumption and number of served requests will be evaluated. Furthermore, the model will be extended to account for user mobility during data transfers.

Acknowledgment

C.F. Chiasserini was supported by NPRP grant # NPRP 5 - 782 - 2 - 322 from the Qatar National Research Fund (a member of Qatar Foundation). M. Gribaudo was supported by ForgeSDK project sponsored by Reply S.R.L. The statements made herein are solely the responsibility of the authors.

References

1. Architecture for mobile data offload over Wi-Fi access networks. `http://www.cisco.com/en/US/solutions/collateral/ns341/ns524/ns673/white_paper_c11-701018.html`, [Online; accessed February-2013]
2. Mobile data offloading for 3g and lte networks. `http://www.alvarion.it/applications/mobile-data-offloading`, [Online; accessed February 2013]
3. Economy + internet trends: Web 2.0 summit. `http://www.morganstanley.com/institutional/techresearch/pdfs/MS_Economy_Internet_Trends_102009_FINAL.pdf` (2009)
4. Cisco Visual Networking Index: Global mobile data traffic forecast update, 2011–2016, Cisco White Paper (feb 2012)
5. Balasubramanian, A., Mahajan, R., Venkataramani, A.: Augmenting mobile 3g using wifi. In: Proceedings of the 8th international conference on Mobile systems, applications, and services. pp. 209–222. MobiSys '10, ACM (2010)
6. Benaim, M., Boudec, J.Y.L.: A class of mean field interaction models for computer and communication systems. Performance Evaluation 65(11-12), 823–838 (2008)
7. Bobbio, A., Gribaudo, M., Telek, M.: Analysis of large scale interacting systems by mean field method. In: 5th International Conference on Quantitative Evaluation of Systems - QEST2008. St. Malo (2008)

8. Cordero, F., Manini, D., Gribaudo, M.: Modeling biological pathways: an object-oriented like methodology based on mean field analysis. In: the Third International Conference on Advanced Engineering Computing and Applications in Sciences(ADVCOM). pp. 193–211. IEEE Computer Society Press (2009)
9. Dimatteo, S., Hui, P., Han, B., Li, V.: Cellular traffic offloading through wifi networks. In: Mobile Adhoc and Sensor Systems (MASS), 2011 IEEE 8th International Conference on. pp. 192 –201 (oct 2011)
10. Gribaudo, M., Cerotti, D., Bobbio, A.: Analysis of on-off policies in sensor networks using interacting markovian agents. In: 4th International Workshop on Sensor Networks and Systems for Pervasive Computing - PerSens 2008. Hong Kong (2008)
11. Guenther, M., Bradley, J.: Higher moment analysis of a spatial stochastic process algebra. In: Thomas, N. (ed.) Computer Performance Engineering, Lecture Notes in Computer Science, vol. 6977, pp. 87–101. Springer Berlin Heidelberg (2011), http://dx.doi.org/10.1007/978-3-642-24749-1_8
12. Hadaller, D., Keshav, S., Brecht, T., Agarwal, S.: Vehicular opportunistic communication under the microscope. In: Proceedings of the 5th international conference on Mobile systems, applications and services. pp. 206–219. MobiSys '07, ACM (2007)
13. Han, B., Hui, P., Kumar, V., Marathe, M., Shao, J., Srinivasan, A.: Mobile data offloading through opportunistic communications and social participation. Mobile Computing, IEEE Transactions on 11(5), 821 –834 (may 2012)
14. Heusse, M., Rousseau, F., Berger-Sabbatel, G., Duda, A.: Performance anomaly of 802.11b. In: INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies. vol. 2, pp. 836 – 843 vol.2 (march-3 april 2003)
15. Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H., Madden, S.: Cartel: a distributed mobile sensor computing system. In: Proceedings of the 4th international conference on Embedded networked sensor systems. pp. 125–138. SenSys '06, ACM (2006)
16. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press, New York, NY, USA, 3 edn. (2007)